

Per capita income and productivity as indices of development and welfare: some comments on Kuznetsians economic history

Tilly, Richard H.

Veröffentlichungsversion / Published Version
Konferenzbeitrag / conference paper

Empfohlene Zitierung / Suggested Citation:

Tilly, R. H. (1983). Per capita income and productivity as indices of development and welfare: some comments on Kuznetsians economic history. In R. Fremdling, & P. K. O'Brien (Eds.), *Productivity in the economies of Europe* (pp. 30-56). Stuttgart: Klett-Cotta. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-329313>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Per Capita Income and Productivity as Indices of Development and Welfare. Some Comments on Kuznetsian Economic History

1. *The Growth Paradigm*

The point of departure of this paper is that a close connection exists between the use of the national income accounts in economic history and the importance for that discipline of what one might call the "Growth paradigm".¹ Insofar as the long run growth of the Wealth of Nations is the concern of economic historians, there is no better frame of reference available for their work than the income accounts. Indeed, I doubt whether meaningful research into the comparative history of economic growth can be done without reference to those accounts (or to some surrogate based on the same principles). Nevertheless, this perspective has limitations—of which two are worth mentioning here. First, it imposes a modern set of values on the past. In Kuznets' words, "the accepted definitions and measures of national product reflect the broad features of modern societies dominated by the ideas of secularism, egalitarianism, and nationalism". These imply that "if we want to contrast modern economic growth with earlier periods and patterns of growth, we must evaluate and appraise the earlier periods also in modern terms in full knowledge that part of the difference would be due to the fact that societies of the earlier times did not share many of the notions of means, ends, and values that constitute impulses to growth in modern times."² The cost of this perspective is our inability to focus on the older values, institutions and activities which may have had to be transformed or eliminated before modern economic growth could begin. Second, the growth paradigm and its accounting complement (national income) implies the primacy of consumption of goods and services as the aim of economic activity and subordinates all other processes—of government, capital accumulation, or even production—to that end. Economic history of this genre is a drama featuring man's conquest of nature for man's material enjoyment. It is a fascinating drama and well worth our attention. However, there are plausible alternatives—for example, the Marxist drama featuring class conflicts and

1. Cf. Tilly, R., *Das Wachstumsparadigma und die europäische Industrialisierungsgeschichte*, in: *Geschichte und Gesellschaft*, 3 (1977); also Parker, W., *Economic History seen through the Income Accounts*, in: *Zeitschrift für die gesamte Staatswissenschaft*, 124 (1968). This volume was also a Festschrift for Walther Hoffmann edited by Giersch H., and Sauermann H., (Quantitative Aspekte der Wirtschaftsgeschichte).

2. Kuznets, S., *Modern Economic Growth. Rate, Structure and Spread*, New Haven 1966.

treating economic growth as a largely unintended consequence of those conflicts.³ For those of us who opt for the growth paradigm, of course, there is need for neither self-congratulation nor apology, only for recognition of a conscious choice and, if possible, acceptance of its limiting implications.

2. An Analogy

The argument of much of the paper is based on a general behavioral assumption and an analogy. The behavioral assumption is that people generally act as if they would rather be rich than poor. The analogy is between individuals and economies. Just as I believe that individuals prefer wealth to poverty, so too do I believe that poor countries strive to become rich ones. And in both cases I believe the relationship to be non-reversible. Rich people and countries do not strive to become poor ones. This non-symmetrical relationship is important for the rest of the argument of the paper, for it serves as a justification for comparing rich and poor countries using the standards (or price weights) of the richer country as a measuring rod. Comparing conditions in this manner is to state how far along a given poor country is on its way to attaining the position occupied by a richer one. And to complete the analogy, we base our comparisons of rich and poor countries on the standard of *per capita* income—which means that we work with the individualistic notion of the representative consumer and make national economic welfare a function of individual welfare.⁴ Countries—or regions—are thus seen as discrete bundles of individuals, a decisive number of which are striving for higher incomes. Were this not the case, the long debate on economic growth of the past three or four decades, it seems to me, would make little sense.

3. Per Capita Income as Welfare Index

Using per capita income as a comparative index of economic welfare implies, then, a unity of opinion about the individualistic ends and the means of economic activity. Kuznets has written: "There is, after all, a strong element of community of human wants and needs, translatable in the modern economic epoch into a set of widely prevalent notions of means, ends, and values of economic activity".⁵ The sad truth, however, is that such unity of opinion in societies over time and space is extremely hard, if not possible, to document empirically. Significant criticism of per capita income as

3. In this sense W. W. Rostow's *The Stages of Economic Growth*, Cambridge 1960, was, in fact, a kind of non-Communist Manifesto, for it did see consumption and technology as the master processes of economic history, if not history *in totum*.

4. On an empirical level, per capita income appears to be much more closely related to a number of important aggregate structural features of developing economies—e. g. the share of total employment and product originating in the agricultural sector, the share of total income spent on foodstuffs and the share of income saved—than is total income. That makes per capita income a more useful instrument for comparative analysis. See esp. Chenery, H., *Structural Change and Development Policy*, Oxford 1979 or Chenery H., and Sirquin M., *Patterns of Development*, Oxford 1975. On the analogy between individual and national per capita income see also Usher, D., *The Measurement of Economic Growth*, Oxford 1980.

5. Kuznets, *Modern Economic Growth*, p. 24.

a welfare index stems, ultimately, from doubts concerning that unity. For the purposes of this paper, this criticism may be divided into five points:

- (1) the problem of non-material or non-economic welfare;
- (2) the exclusion of non-market activities;
- (3) the definition of final (or intermediate) goods;
- (4) the assumption of constancy of preferences and production possibilities; and
- (5) the problem of income distribution and community welfare.

a) Non-material Welfare

The first criticism is that improvements in economic welfare which rising per capita incomes could conceivably reflect say nothing about non-material or non-economic welfare. Non-materialist ends might have priority over material ones in certain societies and, theoretically, satisfaction of the former could deteriorate as a result of improvements in respect to the latter.⁶ This possibility is discussed below in connection with "social indicators", but in the absence of clear evidence to the contrary, we can only assume that changes in economic welfare are not systematically (and negatively) related to changes in non-material wellbeing.

b) Non-market Activities

The second and third criticisms really turn on the proper definitions of the ends of economic activity and the resultant definition of the final products going into national income calculations. National income is a flow statistic reflecting market transactions over a given period. For some countries and periods we have imputations for the value of non-marketed goods and services such as rental income from owner-occupied housing and farm-consumed agricultural products, but, on the whole, market transactions are disproportionately represented. This means that, on the one hand, comparisons of per capita income between developed market economies and societies in which specialization of economic activity and 'hence' market relationships are only weakly developed, could be biased against the latter unless corrections are made for their relatively significant non-market activity. On the other hand, there are some offsetting biases against modern economies for which imputations might well be in order, e.g., the productive work of housewives (as substitutes for domestic servants), time spent in educational institutions, or the value of leisure time generally. Thus, Nordhaus and Tobin have estimated the value of non-market activity in the U.S. in the 20th century at between 40 and 50 percent of GNP while Kuznets has suggested an upward adjustment for the leisure in the same country of as much as 40 percent of the estimated national product.⁷ My impression is that the bias against underdeveloped countries will be strongest for comparisons covering the transitional or "take-off" phase of industrialization, subsequently turning the other way. For western European countries, I suppose, the shift in bias for intertemporal

6. See Gould, J. D., *Economic Growth in History*, London 1972, pp. 5-6.

7. See Nordhaus W., and Tobin J., Is Growth obsolete? in: Moss M. (Ed.), *The Measurement of Economic and Social Performance*. (Studies in Income and Wealth, vol. 38) N.B.E.R., N.Y. 1973, and Kuznets, *Modern Economic Growth*, pp. 220-34 and esp. p. 221. Also Kendrick J., *Economic Accounts and their Uses*, N.Y. 1972. For a brief discussion of this question as applied to American economic history see. Davis L., et al., *American Economic Growth. An Economist's History of the United States*. N.Y. 1972, pp. 42-50. Cf. also Section 5 below.

comparisons would come around 1900. This is a matter which only further research can clarify.

c) Intermediate and final Products

The distinction between inputs, intermediate and final products lies behind a further set of possible biases. National income, as indicated, is conventionally defined as a net flow of final goods and services over time. To avoid double-counting, the value of products used in the production of other final products must be deducted from the value of total output—as in the classic textbook case of the flour used in the production of bakery goods. Problems arise when goods and services satisfy intermediate and final demands and convention assigns them exclusively to *one* of those two classes, or where such conventions vary across time and countries. Per capita income comparisons are biased upward in favour of the more industrialized countries where goods and services such as vehicles, transportation, water supply, sanitation and policing—which are in part costs of urbanization and industrialization and hence akin to intermediate products—are treated as part of final products. This amounts to double counting insofar as other final products embody these costs.⁸ The bias is easy to conceptualize, but in practice, it is virtually impossible to distinguish the part of the total product which is intermediate from that which represents final consumption. Some urban amenities, after all, do (or could) reflect increased consumer utility. Economic historians working in this area will have to decide for each country and period under investigation, (a) which items are ambiguous and (b) how to allocate them.

A similar difficulty relates to the role of capital formation. Net capital formation is commonly regarded as part of final product—on the convincing grounds that it forms the basis of long-run and future consumption. It is difficult to identify, however, because (a) some activities or commodities can be defined as either capital formation or intermediate product and because (b) the flow or capital goods over time is a gross figure and will include the production of replacements for capital used up over a period, i.e. capital consumption allowances, whereas there is no clear rule for estimating the latter. Inter-country and intertemporal comparisons of per capita income obviously will be biased against those economies which work with the narrowest definition of capital formation and/or make the largest deductions for depreciation. Typical problem topics are the treatment of government expenditure on social overheads or infra-structure as intermediate products, the treatment of current spending by integrated business firms on construction of new plant and equipment as intermediate output, or the maintenance of standard deductions on a capital stock of rising durability.⁹ This should be, I suggest, an important target area for historical work on comparative real incomes.

8. Kuznets, *Modern Economic Growth*, pp. 225–27. Intersocietal and intertemporal comparison will reveal some of these “intermediate” goods to be only present in the more developed economy, thus posing, in addition, a weighting problem. More on this below. See also on all of these problems Ruggles N. and R., *The Design of Economic Accounts*, N. Y. 1970, esp. pp. 38–48. For Germany, Stobbe A., *Volkswirtschaftliche Gesamtrechnung*, in: *Handwörterbuch der Wirtschaftswissenschaft*. Vol. 8, Stuttgart and N. Y. 1980.

9. Capital formation raises problems of theory into which the discussion above does not go. See Usher, *Measurement*, esp. Chapter 5. On the measuring problems also Kuznets, *Modern Economic Growth*, Chapter 5.

d) *The Assumption of Constant Preferences*

Intersocietal comparisons of per capita income levels are dogged by the necessity of the unrealistic assumption of constant preferences and production possibilities. Much scattered evidence exists on the variability and mutability of tastes through history, for example, in the discussion of protoindustrialization or of the economics of peasant society, and who could deny that the emergence of new products and product quality changes are an important part of the history of economic development?¹⁰ Strictly speaking, absolute, incontrovertible proof of vast differences in preferences across countries or time should rule out income comparisons *qua* welfare

Table 1: Illustration of Real Income Measurement over Two Periods under Different Assumptions

Period 1					
	<u>Qo</u>	<u>Po</u>	<u>Qo Po</u>		
Good A	20	2	40		
Good B	50	1	50		
Income			90		
Period 2					
	<u>Q1</u>	<u>P1</u>	<u>Q1 P1</u>	(a) <u>Qo P1</u>	(b) <u>Q1 Po</u>
Good A	40	4	160	80	80
Good B	100	2	200	100	100
Income			360	180	180
Period 3					
	<u>Q2</u>	<u>P2</u>	<u>Q2 P2</u>	(a) <u>Q2 P1</u>	(b) <u>Q1 P2</u>
Good A	160	2	320	640	80
Good B	200	2	400	400	200
Income			720	1040	280
Index I	: $\sum Q1$	P1	$\sum Qo Po$	= 400	
"	IIa : $\sum Q1$	P1	$\sum Qo P1$	= 200	(Paasche Index)
"	IIb : $\sum Q1$	Po	$\sum Qo Po$	= 200	(Laspeyres ")
"	IIIa : $\sum Q2$	P2	$\sum Q1 P2$	= 257	(Paasche ")
"	IIIb : $\sum Q2$	P1	$\sum Q1 P1$	= 289	(Laspeyres ")

10. For the general problem, see Gould, *Economic Growth*, pp. 7-9; for "protoindustrialization" and consumer preferences see. Kriedte P., et al., *Industrialisierung vor der Industrialisierung. Gewerbliche Warenproduktion auf dem Lande in der Formationsperiode des Kapitalismus*, Göttingen 1977, esp. Chapter 2 pp. 138-54; on "peasant economics" see Chayanov A. V., *On the Theory of the Peasant Economy*, Edited by Thorner D., et al., Homewood 1966; Mathias P., has raised almost the same question in "Adam's Burden: Historical diagnoses of poverty", *The Transformation of England*, London 1979.

comparisons. Such proof, however, is not generally available for the period of modern economic growth (since around 1750) and so we follow Kuznets and others in *assuming* the broad community of wants and needs across countries and time that the income comparisons require.¹¹ In so doing, to be sure, we are not free to do as we please and are obligated to be as specific as we can about possible distortions (or biases) in the comparisons executed. In this connection we tread the ground well known to economists as the "index number problem"—the essence of which is the difficulty of comparing magnitudes which cannot be compared. For comparisons of real income—be they intertemporal or intersocietal—are only meaningful insofar as they involve indexes of prices and quantities of goods and services having a common denominator, and choosing the latter invariably involves creation of biases. To make this point clearer, a brief digression on price index comparisons follows here.

For illustration purposes, take the example of a simple economy producing two final commodities A and B and compare two periods. Table 1 below depicts the two situations (with Q representing quantity purchased and P the price per unit of commodity). Three important, if banal, conclusions can be drawn from the illustration. First, comparing the products of price and quantity of different periods is not a meaningful comparison of welfare if prices *and* quantities change, for price changes alone do not represent changes in well-being and must be eliminated by deflation. This we may do by multiplying the quantities in both periods by the prices of period 1 (Laspeyres Index) or of period 2 (Paasche Index). Either of these exercises will produce the required common denominator and desired *real* income comparison.¹² Second, the choice of deflator, i.e., the period prices used as common denominator, will have no effect on the welfare comparison only if either quantities of both goods or prices of both goods change at the same rate as between two periods (as between period 1 and 2 in our table). Third, if relative prices and quantities change (and that is what we speak of when changes are not equiproportional) and quantities are rising, then the choice of first period prices will generally produce, *ceteris paribus*, a higher rate of change between periods than the use of end-period prices (in the comparison between period 2 and 3 of our example, e.g. 189 compared to 157 percent). This makes sense in terms of the theory of demand, for that postulates a generally negative relationship between price and quantity demanded and using the higher first period prices to value the larger end-period quantities sets aside that "law of demand" for the end period, so to speak, thus "permitting" consumers to buy as much at higher relative prices as they did when those prices were lower. Just the opposite applies to the use of end-period prices as weights (These yield a lower rate of real in-

11. J. Mokyr's brief survey of demand as a factor in the Industrial Revolution does not explore the possibility of preference shifts except in connection with a presumed trade-off between leisure and money income. He concludes that autonomous demand shifts were of little demonstrable importance for industrialization and pleads for supply-oriented analysis. If his survey is representative, there are few data available on this question. See Mokyr J., *Demand vs. Supply in the Industrial Revolution*, in: *Journal of Economic History*, 37 (1977). Cf. also Mathias P., "Leisure and Wages in Theory and Practice, *The Transformation*, who identifies some possible data sources but on the whole concurs with Mokyr's judgement.

12. In terms of our table Index I has no relevance. For completeness' sake it may be pointed out that price indexes bearing the names Laspeyres and Paasche have the opposite constructions, with the Laspeyres = $\Sigma p_1 q_0 / \Sigma p_0 q_0$ and the Paasche = $\Sigma p_1 q_1 / \Sigma p_0 q_1$.

come change than the "law of demand" would seem to warrant). It is in this sense that we may speak of a "bias" in estimates of *real* income based on the deflators or price weights used.

Turning back to the historical problem of real income estimates, we may ask whether the empirical record confirms such theoretical expectations. The answer is rather ambivalent. There is some seeming confirmation, both for intertemporal and cross-country comparisons. It is believed, for example, that the rate of growth of real incomes in the U.S., 1840–1900, is higher when 1860 prices are used as deflators than when 1900 prices are so employed.¹³ And estimates of the GNP growth of the Soviet Union, 1928–37, employing 1928 price weights are nearly 100 percent larger than estimates using 1937 ones (11.9 percent per annum instead of 6.2. percent!). Sectoral time series studies support this argument.¹⁴ On the other hand, there are exceptions, demonstrated, for example, in a careful study of Swedish income growth by Krantz and Nilsson. In part, this reflects differing levels of aggregation: the broader the categories aggregated, the weaker the substitution effect. In any case, the exceptions remind us that even a theoretical discussion of index number bias in growth measurement must make provision for the possibility of demand shifts (caused either by shifts in tastes or income elasticities of demand).¹⁵ In fact, on theoretical grounds alone, with income effects compensating substitution effects, we can expect index number bias to be negligible.

However, no discussion of the historical use of price indexes will be complete without a few words on the conceptually trivial, but practically significant, question of data comparability. In a strict sense, the world of economic theory with its prices and quantities of individual commodities has no counterpart in reality, and both economic historians and national income accountants have to make do with improvisations and analogies. National income statistics reflect average prices and quantities, but it is apparent that such averages reflect both different data processing operations and different types of transactions. Take the average price of a ton of Ruhr coal in 1855 and 1900: the estimated difference of 21 Pfennige (8.53 (1900) and 8.32 (1855)) or two percent, seems small, but it refers to two quite different commodities: anthracite or hard coal in the first year, and bituminous or soft coal in the second.¹⁶ Now

13 See Davis, *American Economic Growth*, p. 49.

14. Gould, *Economic Growth*, pp. 18–20, discusses this point, citing Moorsteen R., and Powell R., *The Soviet Capital Stock, 1928–1962*, Homewood 1966 and also Gerschenkron A., *A Dollar Index of Soviet Machinery Output, 1927–28 to 1937*, Santa Monica 1951. Gerschenkron has elsewhere discussed this phenomenon of a Laspeyres vs. Paasche "bias" as an essential part of growth, so that the phenomenon itself has been dubbed the "Gerschenkron Effect." See Gerschenkron A., *Economic Backwardness in Historical Perspective*, Cambridge 1962, Chapter 8 and 9, where the difference between the two prices indexes is seen as a measure of structural change.

15. Krantz O., and Nilsson C.-A., *Swedish National Product, 1861–1970*, Lund 1975, deal explicitly with the "Gerschenkron Effect" and attribute its virtual absence in Sweden to significant demand shifts (Ibid., 196–202). See also Solow R., and Temin P., *Introduction: The Inputs for Growth*, in: *Cambridge Economic History of Europe*, Vol. 8, Edited by Mathias P., and Postan M., Cambridge 1978, p. 6.

16. Cf. Holtfrerich C.-L., *Quantitative Wirtschaftsgeschichte des Ruhrkohlenbergbaus im 19. Jahrhundert*, Dortmund 1973, p. 18.

coal probably represents a resolvable problem, but what about iron or steel? Here it is not merely a matter of specifications of physical properties of the commodity, e.g., its subdivision into components such as bars, plates and rails and standardized weight measures, but also of knowing whether the prices averaged reflect a standard procedure for estimating transportation cost from plant to representative consumer, discounts for volume and cash purchases, etc. etc.¹⁷

Then there is the question of uncertainties in the estimation of quantity data. For example, for the Prussian and German *agricultural* sector, the choice of period, the treatment of intermediate products, and assumptions about slaughter rates and weights are much more important determinants of the measured rate of growth of output than are the choice of price weights. This can be demonstrated by means of the following estimates:¹⁸

Aggregate Output,	Prussia,	1816-49	(vF) 2,1% p.a
Aggregate Output,	Prussia,	1816-49	(GH) 2,2% p.a
Net Output,	Prussia,	1816-49	(RT) 2,6% p.a
Net Output,	Prussia,	1816-52	(RT) 2,1% p.a
Net Output,	Germany 1846/49-1910/13		(L)=1,46; (P)=1,40 p.a
Net Output,	Germany 1850/54-1910/13		(L)=1,85; (P)=1,83 p.a

Index number problems are not a negligible factor and warrant further consideration in connection with productivity measurement. But the point is, historians of 19th century productivity may face more dangerous enemies.

We have already called attention to the symmetry of time series and cross-sectional comparison. We thus expect the estimated income differences between rich and poor countries to be larger using the latter's prices as weights. Patel's experiments with Indian and American data for 1959 showed a difference of 100 percent, i.e., India's per capita output was more than twice as high in U.S. dollars when measured in U.S.

17. Morgenstern O., *On the Accuracy of Economic Observations*, 2d Ed., Princeton 1965, esp. Chapter 10. It should be pointed out that inferences about costs on the basis of price data depend on assumptions about competition which require investigation. They can be critical, as the discussion of British, German and American productivity in the steel industry seems to indicate. See, e.g. Allen R. C., *International Competition in Iron and Steel, 1850-1913*, in: *Journal of Economic History*, 39 (1979), esp. pp. 933-37, where a debate with the study of McCloskey D., *Economic Maturity and Entrepreneurial Decline*, Cambridge 1974, is joined. See also Webb S., *Tariffs, Cartels, Technology and Growth in the German Steel Industry, 1879 to 1914*, in: *Journal of Economic History*, 40 (1980), esp. pp. 321-23.

18. von Finckenstein, Graf M. W., *Die Entwicklung der Landwirtschaft in Preußen und Deutschland, 1801-1930*, Würzburg 1960, Helling G., *Berechnung eines Index der Agrarproduktion in Deutschland im 19. Jahrhundert*, in: *Jahrbuch für Wirtschaftsgeschichte*, 4 (1965); Tilly R., *Capital Formation in Germany in the Nineteenth Century*, in: Mathias P., and Postan M., (eds.), *Cambridge Economic History of Europe*, vol. 7, Cambridge, 1978; Jacobs A., and Richter H., *Die Großhandelspreise in Deutschland von 1792 bis 1934, Sonderheft des Instituts für Konjunkturforschung*, Berlin 1935. The abbreviations vF, GH and RT in the text refer to the sources von Finckenstein, Gerhard Helling and R. Tilly, respectively. 'L' to Laspeyres price weights and 'P' to Paasche ones.

prices as when measured in Indian prices and much closer to American levels.¹⁹ A less extreme but nevertheless significant difference was produced by Gilbert and Kravis—by somewhat different methods—in their classic study of purchasing power parity exchange rates between European countries and the U.S. in the early postwar period.²⁰ Table 2 summarizes their findings.

Table 2: Income per capita in Different Countries According to Exchange Rates and Purchasing Power Parities, 1950 (U.S. per capita Income = 100)

<u>Country</u>	<u>Exchange Rate</u>	<u>U.S. Prices</u>	<u>European Prices</u>
U.S.A.	100	100	100
U.K.	37	63	53
France	35	53	42
Germany	26	43	33
Italy	16	30	22

Source: M. Gilbert and I. Kravis,
An International Comparison of
National Products and the Purchasing
Power of Currencies (Paris, n.d.)

Last but not least I should mention here O'Brien and Keyder's study comparing Great Britain and France as one of the first significant attempts to extend this kind of analysis to 19th-century economic history.²¹ In this case, however, the fact that French per capita incomes seem relatively higher with French price weights than with British ones is not unambiguously interpretable in terms of our rich country—

19. Patel S., *The Economic Distance Between Nations: Its Origin, Measurement and Outlook*, in: *Economic Journal*, (1964). See also Kuznets, *Modern Economic Growth*, pp. 374–84.

20. Gilbert M., and Kravis I., *An International Comparison of National Products and the Purchasing Power of Currencies. A Study of the U.S., the U.K., France, Germany and Italy*, Paris n.d.

21. O'Brien P., and Keyder C., *Economic Growth in Britain and France, 1780–1914. Two Paths to the 20th Century*, London 1978. It is useful to note, however, that in this comparison, differing national output structures are used to weight two national consumption “baskets” (including items common to both countries) which, by means of substitution of each country's prices into the other country's “basket”, yield two “exchange rates”. These are then applied to the money income of one country to permit one-currency income comparisons. That is, p_0q_0/p_1q_0 or p_1q_1/p_0q_1 where 0 = Great Britain and 1 = France. Their exercise involves only conversion of French incomes into Sterling or: French income/ p_0q_0/p_1q_0 and French income/ p_1q_1/p_0q_1 .

poor country dichotomy; for the country differences may have been small. In discussing how this method gets around some of the difficulties of using official exchange rates for a cross-country income comparison, O'Brien and Keyder suggest calculating *two* rates of exchange for every two-country comparison: the purchasing power parity of sterling in terms of francs (or the number of francs needed to purchase a basket of goods representative of British consumption patterns costing 1 £ in Britain) and the purchasing power parity of francs in terms of sterling. This suggestion correctly emphasizes that (a) the rate reflecting one country's price weights, say, Britain's, permits conceptualizing how well off an average inhabitant of that country would be in Britain with the average income of inhabitants of another country—in this case, France; that (b) the same experiment with the other country's price weights permits the opposite comparison, (c) that both rates are equally "valid", and (d) that the difference between the incomes so converted reflects differences in preference patterns but also gives us an idea of the maximum and minimum size of real income differentials. Like all such explicit comparisons it is an exercise in hypothetical history. This particular case, to be sure, is a double exercise.

These observations return us to Kuznets' interpretation of real income comparisons across time and space as being products of a point of view rather than reflections of universally objective measurements. The point is well taken, but we should not forget that Kuznets also suggested that in long-run historical comparisons, some points of view may be more valid than others. Where income gaps between countries are large, he recommended use of the preference and production patterns, i.e., the price weights, of the more advanced, high-income country for comparative purposes on the grounds that poorer countries strive to become richer but not vice-versa.²²

This brings us full circle and back to the remarks about the universal community of wants and needs and the analogy between rich and poor persons with which this section began. It amounts to an endorsement of the use of per capita income as an index of economic development, though the endorsement is a qualified one. In the next section, we must conclude our discussion of that index by examining what is perhaps the single most important qualification—the unresolved problem of distribution.

e) Income Distribution and Community Welfare

An increase in a country's per capita real income could mean an increase in its economic welfare in the sense of increased satisfaction of material wants, but such an increase will reflect the distribution of income, since, obviously, only those wants backed up by income can be made effective. One could take the position that every society gets the income distribution it deserves and regard per capita income as ever-optimally distributed—be it in a social-democratic, welfare-state economy, a *laissez-faire* liberal one, or a totalitarian communist dictatorship. But this would be panglossian. Alternatively, one can impose modern distributional "welfare functions" on the historical situation investigated. On only moderately egalitarian assumptions about the utility of income to different classes of individuals in society, we have to recognize, it seems to me, that increases in per capita income might not reflect increases in aggregate welfare at all, for example, if they were accompanied by a sharp

22. Kuznets, *Modern Economic Growth*, pp. 23, 484–85.

redistribution of income in favour of the wealthiest members of the community and/or against a great majority of low-income receivers.²³

Given the fact of distributional inequality and the practical impossibility of assigning generally accepted utility weights to different income groups, quite a few scholars have chosen not to interpret real income per capita as an index of welfare at all, but rather as an index of productive capacity, as an index of *potential* welfare, so to speak. According to Harvey Leibenstein, for instance, increasing per capita income represents increasing "possible achievement" i.e., a larger sum available for potential redistribution, should that be found desirable.²⁴ I disagree with the notion of "potential welfare"²⁵ and wish to return to the related interpretation of real income per capita as productive capacity shortly; but for the moment, let us note that such a reaction depends on one's ideas about the behavior of income distribution. In the face of compelling evidence confirming the stability of income distribution across time and countries, most economists and economic historians, I suspect, would find it difficult not to interpret increases in per capita real income as improvements in community welfare. And evidence showing non-negligible increases in the inequality of income distribution, it follows, could be seen as reductions in community welfare, deductible, as it were, from any increases in per capita real income.

In any case, that is the sensible approach followed in a number of important treatises on economic development, notably those by H. Chenery and his collaborators.²⁶ The schemes devised in these studies weight income growth in the different income classes by the number of persons in them. Given the disproportionately large share of population in the lowest groups, this amounts to assigning utility points to increases in the share of income *increases* going to the lowest income groups of a given country. The logic of this procedure derives from its frequently practiced opposite: to view aggregate per capita real income growth as welfare growth is, in fact, to weight increases in the average income of the wealthy, say, the top 20 percent of income re-

23. Usher, *Measurement*, Chapter 3, lists identical tastes, equal shares in ownership of the factors of production and/or unitary income elasticities of demand for all goods as the conditions for interpreting real income estimates based on observed prices and quantities as a community welfare index.

24. Cf., e.g., Leibenstein H., *Economic Backwardness and Economic Growth*, N. Y. 1963, Chapter 2; also Viner J., *International Trade and Economic Development*, Oxford 1953, Chapter 6.

25. The problem with "potential welfare" is that it is misleading, for a situation with more potential welfare can quite easily be a situation with less actual welfare if the contingent redistribution does not take place. The notion of "potential welfare" thus settles nothing. On this and other related matters, see. Sen A., *The Welfare Basis of Real Income Comparisons: A Survey*, in: *Journal of Economic Literature*, 17 (1979). Sen, in fact, proposes some measures of inequality of income distribution which are worth considering, but he appears, in general, to take the position that a country's economic welfare is not measurable in terms of its per capita income. I have a less rigorous understanding of economic welfare than Sen and persist, in this paper, in associating it with per capita income—subject to one qualification to be mentioned shortly. See also Usher, D., *The Welfare Basis of Real Income Comparisons: A Comment*, in: *Journal of Economic Literature*, 18 (1980).

26. Chenery, H., et al., *Redistribution with Growth*, Oxford 1974; Chenery H., *Structural Change and Development Policy*, Oxford 1979; Chenery H., *Armut und Fortschritt—Alternative für die Dritte Welt*, in: *Finanzierung und Entwicklung*, 17 (1980); also Sen, *Welfare Basis*, pp. 30–31, and some of the literature cited there.

ceivers—who typically obtain 50 percent of a given income increase in poor countries—about 10 times higher than the income gains of the representative poor—whose aggregate increment typically accounts for five percent of the total.²⁷ Alternatively, weights can be assigned to increases in the share of the population living above some materially defined standard of poverty. The importance of such possible adjustments lies in the fact that growth of per capita incomes has not automatically contributed to alleviation of poverty in poor countries in recent years. Indeed, according to Chenery, in some places and times income growth has achieved less than specific distributional policies have done. In this connection we are invited to compare the experience of slow growers such as Cuba or Sri Lanka with fast growers such as Brazil.²⁸ Historical extensions of the argument readily suggest themselves. They run from the famous “Standard of living” debate concerning British workers during the Industrial Revolution, through S. Kuznets’ well-known thesis on the inverted U-curve of income inequality during economic development (e.g., increasing inequality in the early stages) to more recent work on Britain and the United States by J. Williamson, P. Lindert and others.²⁹ I have no wish to review this literature here and only mention it as a way of suggesting that the distribution of income may represent an important modification of per capita as a long-run development welfare index.

However, the word “may” in the previous sentence was used advisedly, for certain problems emerge with this use of distributional considerations that have not yet been satisfactorily resolved. Their mention therefore concludes this section of the paper. The relevance of distribution for welfare interpretations of per capita income growth will depend on answers to three questions: (1) to what extent are we free or obligated to impose our presumably modern welfare standards on the past—even in the face of evidence on the prevalence of wholly different welfare notions among the populations being investigated. Put in a comparative context: do we impose one standard on two societies and will there be “bias” as a consequence? (2) To what extent does a distributional correction of per capita income indexes imply recourse to a standard of individual aspirations which, strictly speaking, requires additional correction, e.g., for average age and life cycle experience of the population? (3) The evidence cited above to the contrary notwithstanding, is a distributional correction necessary? Or rather: what is the long-run relationship between income growth and the equality of its distribution? I submit that if our answer to the last question is “positive”³⁰ we can

27. Chenery suggests (in *Redistribution with Growth*) the following measure of development as welfare: $G = w_1g_1 + w_2g_2 + w_3g_3 + w_4g_4 + w_5g_5$ where g = the mean income of each quintile of the population of income recipients and w = the population weight of each quintile.

28. Cf. esp. Chenery, *Armut und Fortschritt*, p. 13; also Chenery et al., *Redistribution*.

29. Cf. Taylor A., (Ed.), *The Standard of Living in Britain in the Industrial Revolution*, London 1975; Kuznets S., *Economic Growth and Income Inequality*, American Economic Review, 45 (1955); J. Williamson, *Earnings Inequality in Nineteenth-Century Britain*, in: Journal of Economic History, 40 (1980); Williamson J., *The Sources of American Inequality, 1896–1948*, in: Review of Economics and Statistics, 58 (1976); and the interesting observations on links between inequality and cost of living indexes by David P., and Solar P., A Bicentenary Contribution to the History of the Cost of Living in America, in: Uselding P., (Ed.), Research in Economic History, Greenwich 1977.

30. Some evidence points in this direction. See e.g. Chenery, *Structural Change*, Chapter 8.

avoid some very difficult conceptual and empirical problems, though that will not obviate the necessity of mobilizing the distribution data themselves. They remain as significant desiderata of the comparative history of income growth.

4. *Social Indicators and the Income Concept*

Dissatisfaction with per capita income as a development and welfare measure produced in the 1960's the so-called "social indicators movement". It represents the search for quantitative indicators reflecting dimensions of social experience presumed not to be covered by the national income accounts. Before taking up the social indicators, however, I would like to make a few comments on some extensions of the concept of national income which are closer to its original meaning and were mentioned briefly earlier under the heading of "imputations". In the early 1970's, W. Nordhaus and J. Tobin offered one of the most comprehensive suggestion for extensions in the form of a "Measure of Economic Welfare" (M.E.W.). As the name suggests, their concern was with developing a more consistent measure of *welfare* than GNP or income per capita were believed to provide. Welfare is defined as the consumption of final goods and services including an allowance for leisure and a deduction for environmental deterioration *plus* the investment expenditures necessary to insure maintenance of the current rate of productivity growth into the future. Their calculation thus (1) divides government expenditures into final consumption and intermediate goods and services ("regrettables" such as defense and other "inputs" to other sectors); (2) divides household expenditures into consumption and investment and intermediate activities (such as expenditures for travelling to work); (3) expands the concept of capital formation to take account of consumer durables, education and health investment, as well as investment (already mentioned) needed to insure a given current rate of productivity advance; (4) expands the concept of final product to assign to leisure time activities consumption values (this is by far the largest modification, amounting to more than GNP for 1929, e.g.); and (5) makes a calculation of the social costs of growth especially, those related to environmental deterioration.

Table 3 suggests their 20th-century importance.³¹

I must confess to mixed feelings about suggestions such as these. On the one hand, we should think seriously about expanding historical national income statistics in a similar direction, for a good *a priori* case can be made for some of the extensions suggested and, in any case, we are faced with the need to estimate the value of historical transactions on the basis of evidence just as suspect and as indirect as that employed by modern "imputers".³² On the other hand, a case can also be made—for some purposes such as the analysis of cycles—for reducing or limiting the measured product

31. This discussion is based on Nordhaus and Tobin, cited in note 6 and the discussion in that NBER volume, see also Usher, *Measurement Journal of Economic History*, Chapter 7, Kuznets, *Modern Economic Growth*, pp. 220-34, and Kendrick J., *Economic Accounts and their Uses*.

32. For example, agricultural income is derived from crop and animal output estimates applying market prices to *non*-market production, and much of non-agricultural income also must be imputed, e.g., that of self-employed artisans.

*Table 3: Gross National Product and M.E.W. in the U.S.A.,
1929-65 (billions of dollars, 1958 prices)*

	1929	1935	1954	1965
1. Gross national product	203.6	169.5	407.0	617.8
2. Capital consumption NIPA	-20.0	-20.0	-32.5	-54.7
3. Net national product NIPA	183.6	149.5	374.5	563.1
4. NIPA final output reclassified as regrettables and intermediates				
a. Government	-6.7	-7.4	-57.8	-63.2
b. Private	-10.3	-9.2	-16.4	-30.9
5. Imputations for items not included in NIPA				
a. Leisure	339.5	401.3	533.2	626.9
b. Nonmarket activity	85.7	109.2	211.5	295.4
c. Disamenities	-12.5	-14.1	-24.3	-34.6
d. Services of public and private capital	29.7	24.2	48.9	78.9
6. Additional capital consumption	-19.3	-33.4	-35.2	-92.7
7. Growth requirement	-46.1	-46.7	-63.1	-101.8
8. Sustainable MEW	543.6	573.4	961.3	1,241.1

NIPA= national income and product accounts.

Source: Nordhaus and Tobin as cited in note 7.

to market or quasi-market transactions. This is owing to the absence of annual data for imputed items and the resultant necessity of extrapolating and interpolating for missing observations. If the trade-off between observable market activities and imputed non-market ones is subject to cyclical influences—modifications by means of trend relationships may produce a distorted view of economic growth patterns and also of welfare, as for example Nordhaus and Tobin's data do for the 1929-35 period of U.S. economic history (when much *involuntary* leisure time emerged). The point is, one's particular research interest may have to determine one's choice of definitions of national product.

Turning away from national income to the "social indicators movement", we should first note that in one sense this movement represents a step backwards: none

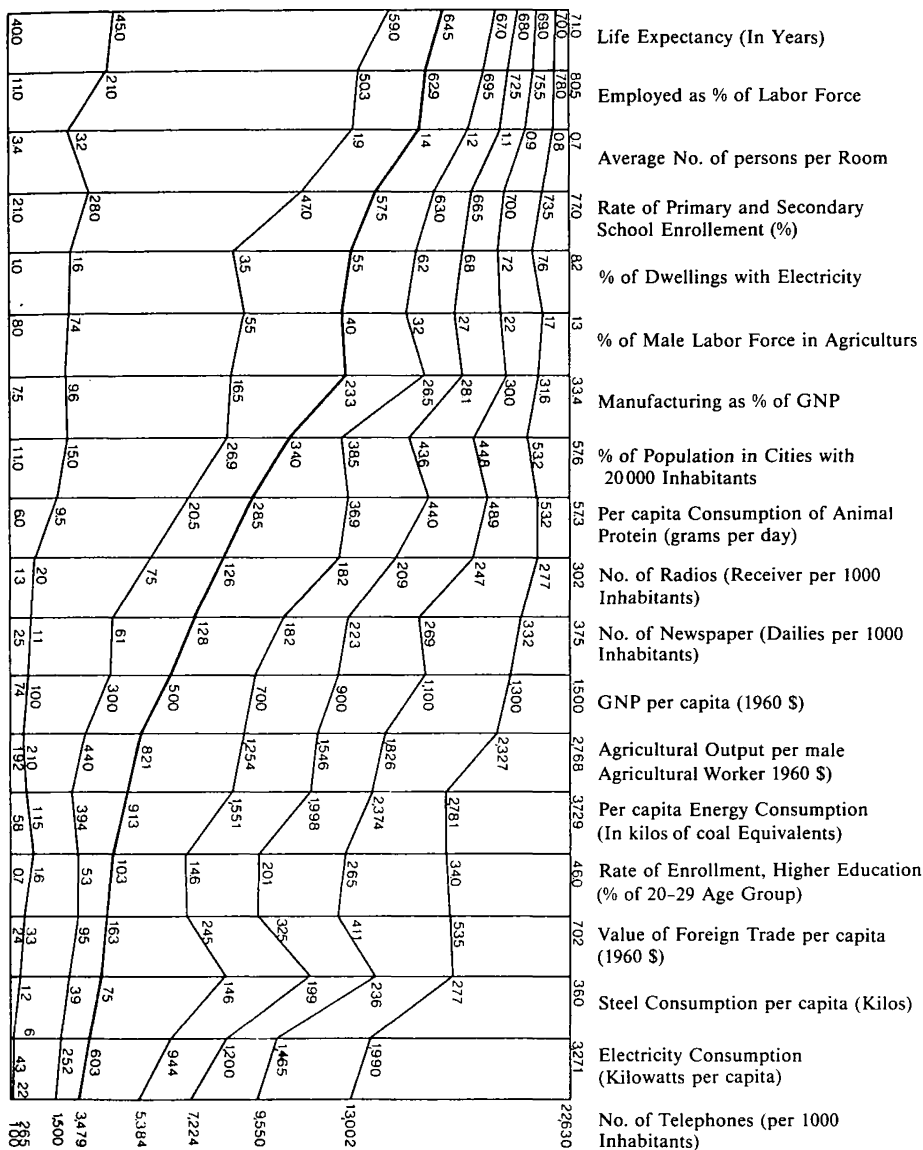
of the proposed indicators are themselves as comprehensive as the income accounts, and in contrast to the components of those accounts they cannot be added up to something comparable to GNP. However, for many of the experts working in this area of applied statistics the data requirements of GNP and its underlying assumptions seemed to go well beyond available knowledge, particularly where comparisons with less developed countries of the past and present were sought. One idea was to collect those concrete data which were believed to represent development indicators, were readily available, and posed no difficult valuation problems: for example, tons of wheat or steel production per annum, numbers of bicycles or radio sets per head of population, number of crude births per 1,000 inhabitants, and so on. Comparison, it was hoped, could thus be extended to countries having no national income statistics.³³ On the whole, however, this attempt has not been particularly successful, either because the indicators or their averages, taken by themselves, had no clear meaning, (either as welfare or capacity indicators) or because, where they were linked via correlation analysis to GNP, they became no more than rather poor proxies for the latter.

Table 4 illustrates the bind we are in. There is good correspondence in an ordinal ranking sense, but the scales of the variables are multi-dimensional and hence, quantitatively non-comparable. This means that for ranking purposes the indicators are superfluous—since we already have per capita income statistics—but for quantitative extensions into times and countries with deficient data, inadequate, especially when we note the irrelevance of some of the indicators for historical work on the 19th century (e. g., radios or electricity).

The “social indicators movement”, however, aimed in two other directions as well. One of those reflects modern-day concern in industrial nations with the “costs of economic growth”. It is in search of indicators of social and environmental change generally believed to affect social welfare. Here, the idea has been to develop an index of the “quality of life” which could be combined with GNP to help to decide whether economic change over a given period has been, on balance, socially beneficial or detrimental for the population affected. This effort is closely related to the extensions of the concept of real income discussed earlier (as “imputations”). It is much too early to say what will come out of this attempt, but two tentative observations seem relevant for our purposes. First, as in the correspondence “test” just displayed, most of the indicators of social well-being are positively associated with changes in GNP per capita. Given the widespread belief that the direction (or sign)

33. See on this, Gould, *Economic Growth*, pp. 11–14; Beckerman W., and Bacon R., *International Comparisons of Income Levels: A Suggested New Measure*, in: *Economic Journal*, 76 (1966); or the discussion of social and economic indicators developed by the UN in Nohlen D., and Nuscheler F., (Eds.), *Handbuch der Dritten Welt*, I, Hamburg 1974. It may be added that Gerschenkron's wellknown approach to European industrialization limited quantitative growth analysis to industrial production, for want of more comprehensive, yet reliable, data. Cf. Gerschenkron A., *The Approach of European Industrialization: A Postscript*, and *Problems of Measuring Long-Term Growth in Income and Wealth*, in: *Economic Backwardness*, esp. pp. 353–54; also Gerschenkron, *The Early Phase of Industrialization in Russia: Afterthoughts and Counterthoughts*, in Rostow W. W., (Ed.), *The Economics of Take-Off into Sustained Growth*, N. Y. 1963, esp. pp. 161–63.

Table 4: (GNP per capita and Corresponding Indicators)



Source: U.N. Research Institute for Social Development (Reprinted in D. Nohlen and F. Nuscheler (Eds.), *Handbuch der Dritten Welt*, I (Hamburg, 1974), p. 248.

of this correlational relationship is negative, that is an important result.³⁴ Second, the indicators are aggregative and say nothing about distributions and/or social inequality. If we wish to correct GNP for changes in the degree of economic inequality, an analogous correction for these indicators will also be in order.

A second prong of the "social indicators movement" has focused on the under-developed countries. Its proponents have interpreted economic growth as part of a broader cultural process embracing social and political change. Such an interpretation called for (a) data on social and political change and (b) analysis of the links between such data and more conventional indicators of economic development. I. Adelman and C. T. Morris, in one of the more ambitious examples of this genre, have shown that the interdisciplinary data approach can enlighten.³⁵ Applying factor analysis to a cross-country array of many economic, social and political indicators—factor analysis being particularly useful where an inductive approach is preferred and no specific hypotheses are used to pre-structure the data—they find that (estimated) GNP per capita is closely associated with the non-economic factors of development, but that the complex "factor" of "socio-economic" indicators (including such variables as "extent of dualism", importance of an indigenous middle class, etc.) corresponds much better to what we commonly think of as the conditions of underdevelopment than the purely economic "factor" does.³⁶ In addition, tracing the links between economic, social and political variables leads Adelman and Morris to a useful three-stage version of underdevelopment in which GNP per capita becomes more closely related to purely economic factors as development proceeds (from "very underdeveloped" to "less developed") and also, in the third stage of underdevelopment, more closely associated with political ones.³⁷ To sum up: these results suggest to me that historians of development have something to gain by an extension of the indices of measurement to non-economic factors, particularly for study of the earliest phases of industrialization. However, for investigations of western European development since around 1850—corresponding to the third phase—they are not likely to be misled very much by relying on per capita income—so long as their main concern

34. Cf. e.g., King M. A., *Economic Growth and Social Development*, in: Review of Income and Wealth, 20 (1974). I owe this reference to Rolf Dumke. For Germany, see Zapf W., *Lebensqualität in der Bundesrepublik. Methode der Messung und erste Ergebnisse*, in: Soziale Welt (1977).

35. Adelman I., and Morris C. T., *Society, Politics and Economic Development: A Quantitative Approach*, Baltimore 1967. Their goal, to be sure, is an improved explanation of economic development, not an analysis of "modernization". That branch of the indicators movement is not discussed here at all.

36. One way of putting this is that GNP per capita categories bring together much more diverse collections of countries with respect to socio-economic characteristics (such as size of traditional agriculture, extent of dualism, etc.) than the "factor" "Socioeconomic Variables" does. See Adelman and Morris, *Society*, p. 169, and the appendix to this paper where some of their results are reproduced. See also Kuznets, *Modern Economic Growth*, pp. 437–60, for a discussion of non-economic characteristics of underdeveloped countries.

37. The Adelman and Morris discussion of this third stage with its emphasis on the emergence of strong leadership commitment to economic growth and corresponding economic and financial policies gives it a resemblance to the experience of some European countries in the 19th century, e.g., Germany in its "Take-Off" phase (in the 1850's and 1860's).

is with *economic* development, and not the social and political changes which accompanied it.

5. Productivity

Aggregate labor productivity bears a close resemblance to per capita income. Indeed, where the entire population is gainfully employed, capital consumption negligible, and the foreign accounts are in balance, the two are virtually identical. This conceptual likeness is important, for it reflects a duality in the way we look at the economy and the real income it generates: it may be seen as a system of prices and quantities of goods and services generating real income and welfare as a function of consumer preferences; or it can be interpreted as a system of production possibilities generating output (real income) as a function of technology. For some purposes either perspective will do; but where interest centers on the measurement of real income as an index of development economists differ on this issue. As suggested earlier, quite a few economists believe that interpreting output or income per capita as an index of productive capacity avoids the problems of interpersonal comparisons of utility and distributional considerations associated with income seen as welfare.³⁸ They see development as an economy's progression from, say, position A to position B, where B represents an economic state capable of producing all of the goods produced (per capita) in position A plus some non-negligible quantity of goods reflecting income growth, call it P—for increased productive capacity. Thus $B = A + P$. The words "capable" and "capacity" are used to stress that while B represents more "potential welfare" than A, it need not mean more realized welfare.

I disagree with this interpretation. In terms of the example just given, I suspect that we are likely to say that an economy is better off in position B than in position A, a statement which to my way of thinking has welfare connotations. The notion of "potential welfare" is, on this view, unnecessary baggage. More importantly, historically relevant comparisons will typically involve economies with different bundles of goods and services. Referring back to our example, we may wish to view states A and B as two separate economies. If we observe them to produce the same products, x and y, but B to produce more of each, there is no ranking problem. However, it is thinkable that we might observe that economy B produces more of final goods x and y than economy A but much less of an intermediate good z (say, transport), mainly because consumers in economy B have localized tastes causing localized consumption and reduced need of transportation. Even if relative prices of goods x and y were identical as between economies A and B, so long as transportation is an intermediate good (and resources are perfectly convertible) there would be a difference in per capita real income between A and B dependent upon demand differences.³⁹

38. Leibenstein, *Economic Backwardness*, p. 12. This distinction has sometimes been applied in the national accounts to the difference between national income at market prices (welfare measure) and at factor costs (productive capacity measure). See, e.g., Gilbert and Kravis, *International Comparison* (cited in note 20). For critical discussion of this distinction see Usher, *Measurement*, Chapter 4, and Sen, *Welfare Basis*.

39. The example and much of this entire discussion is derived from Usher, *Measurement*, Chapter 4.

Alternatively, turning to the topic of productivity comparison, we can imagine economy A having a higher productivity or output per head than B in the production of the only commodity it is capable of producing, say, good x, but B having the ability to produce a greater variety of goods and services, say, goods x and y rather than just x. Any comparison of the productivity of the two economies will have to take account of a bias connected with the comparison of non-identical bundles of goods and services. We have here, once again, the “index number problem”, in this case in an extreme form. Now, if we were to substitute into our example the economies of the American ante-bellum South (for B) and the North (for A), we could cite in support of our argument an important criticism which P. David and P. Temin levelled against Fogel and Engerman’s analysis of slavery in their book on the American South—*Time on the Cross*.⁴⁰ Those reviewers argued that the extent of the South’s productivity (or “efficiency”) advantage over the North in agriculture was dependent on the relative price of cotton (and other goods) in 1860 and not only or even mainly on the superior productive organization and “labor efficiency” of Southern plantations. Because the South could produce much cotton and because world demand for it was buoyant, the value of its agricultural output could and did expand above Northern levels. These values, divided by estimates of the available productive factors, produced a “relative efficiency” or “productivity” differential favoring the South. But behind that differential lay the crucial weight of demand. For this reason, David and Temin suggested replacing the terminology “relative efficiency”—having physical and technical connotations—with the label “revenue—getting efficiency”.⁴¹

I have neither the competence nor the desire to adjudicate in the ongoing debate on American slavery, but the particular criticism of interpretation of productivity in-

40. The literature of criticism of this book is immense. Some of the landmarks are: Fogel R. W., and Engerman S. L., *Time on the Cross*, Vol. I: *The Economics of American Negro Slavery*; Vol. II: *Evidence and Methods*, Boston 1974; David P., and Temin P., *Slavery: The Progressive Institution?*, in: *Journal of Economic History*, 34 (1974); the entire issue of *Explorations in Economic History*, 12 (1975); R. W. Fogel and S. L. Engerman, *Explaining the Relative Efficiency of Slave Agriculture in the Antebellum South*, in: *American Economic Review*, 67 (1977); Haskell, T., et al., in: *American Economic Review* 69, (1979); and Fogel R. W., and Engerman S. L., *Explaining the Relative Efficiency of Slave Agriculture in the Antebellum South: Reply*, in: *American Economic Review*, 70 (1980); and David P., et al., *Reckoning with Slavery*, N. Y. 1976.

41. The David-Temin discussion (in 1974 and 1977) suggests that the bias favoring the South is unknown, but that it could, in the extreme, explain the whole of the relative advantage recorded by Fogel and Engerman. Interested readers are referred to the Fogel-Engerman reply to this criticism in the *American Economic Review*, 70 (1980). I should add that the point I am making is not identical with the general thrust of the David-Temin critique of *Time on the Cross*. For the purposes of the present argument, the South’s alleged productivity lead can be seen as a welfare advantage over the North, though one which is demand-dependent. David and Temin, however, would doubt whether the assumptions of welfare economics apply to unfree societies such as the antebellum South at all. I find it hard to resist their arguments on this point, though I have the feeling that a strict construction of their remarks could considerably narrow the range of comparative economic growth history. My point, it must be reemphasized, is that productivity comparisons will involve welfare judgements in many relevant cases. The reader is referred, once more, to the rather loose, non-optimal concept of welfare applied here—and discussed in sections 3d. and 3e. of the paper.

dices as indicators of physical or technical efficiency articulated there does have general validity for all comparative studies of productivity. For there are not very many sets of productivity figures which will stand comparison without translation into some common denominator. There are some, e.g. grain yields, or tons of coal mined per miner-hour, but their range of application is limited. The desired translation, in any case, will involve consideration of prices and quantities of commodities, weighting them, and as an inevitable part of the counterfactual experiment which choosing such weights involves, bias. Historians of comparative productivity, that is the conclusion to be drawn here, should neither deny the existence of such biases nor seek to escape them, but should instead attempt to construct their studies so that the inevitable biases will not invalidate their results.

Although I do not believe that productivity data offer us more "objective" evidence of economic progress than do those on real incomes per capita, they are nevertheless essential. They are essential, because the supply side is just as indispensable to analysis of economic development as the demand side and probably more accessible to research. For this reason an entire generation of economists and economic historians over the past several decades has discussed economic growth in terms of them. Those scholars have conceived of economic growth (or development), for better or worse, as a technical process in which inputs of productive factors are transformed via a production function into output. They have retained this schema of relationships while disaggregating the process down to the sectoral or regional level, and they have managed to refine some of the input measures (e.g., by converting the labor input with the help of educational data into a range of labor inputs of differing quality) without making the entire set of accounts—for that is what this system of production—function relationships is—inconsistent or noncomparable.⁴² Given its widespread acceptance, this system of relationships offers possibilities for further comparative study which it would be foolish to ignore.

A possibly superficial, but nonetheless logical, argument for focusing on productivity change is the latter's contribution to long-run changes in per capita income. Long-run economic growth in those countries for which estimates exist, has resulted largely from increased output per unit of input, i.e., from productivity growth.⁴³

42. This is clearly not the proper place for a bibliographical survey, but a few major contributions to the "growth paradigm" should be mentioned: Abramovitz M., *Resource and Output Trends in the United States since 1870*, in: *American Economic Review*, 56 (1956); Solow R. M., *Technical Change and the Aggregate Production Function*, in: *Review of Economics and Statistics*, 39 (1957); Denison E., *The Sources of Economic Growth and the Alternatives before Us*, N.Y. 1962; Ibid, *Why Growth Rates Differ*, Washington 1967; Jorgenson D., and Griliches Z., *The Explanation of Productivity Change*, in: *Review of Economic Studies* (1967); Kuznets, in *Modern Economic Growth* and elsewhere has extended the empirical basis for growth studies and thus deserves mention here. I have found useful discussions of growth and technical change in Usher, *Measurement*, Chapter 12, and also in Gould, *Economic Growth*, Chapter 5.

43. See, for example, the discussion of Denison's findings for the U.S. (1909–57) in Kuznets *Modern Economic Growth*, pp. 80–82 (where an annual rate of growth per capita income of 1.44 per cent is attributed to the growth of capital and land inputs (.18), quality improvements of labor (.58) and factor productivity (.67). See also the correlation between per capita income and output per man-hour for 14 industrial countries (1870–1965) indicated in Gould,

There is no need to reproduce the basic estimates here, for they are common knowledge. Rather more useful may be some discussion of the qualifications of those findings, for they are likely to be relevant to some of the historical studies of productivity currently getting underway.

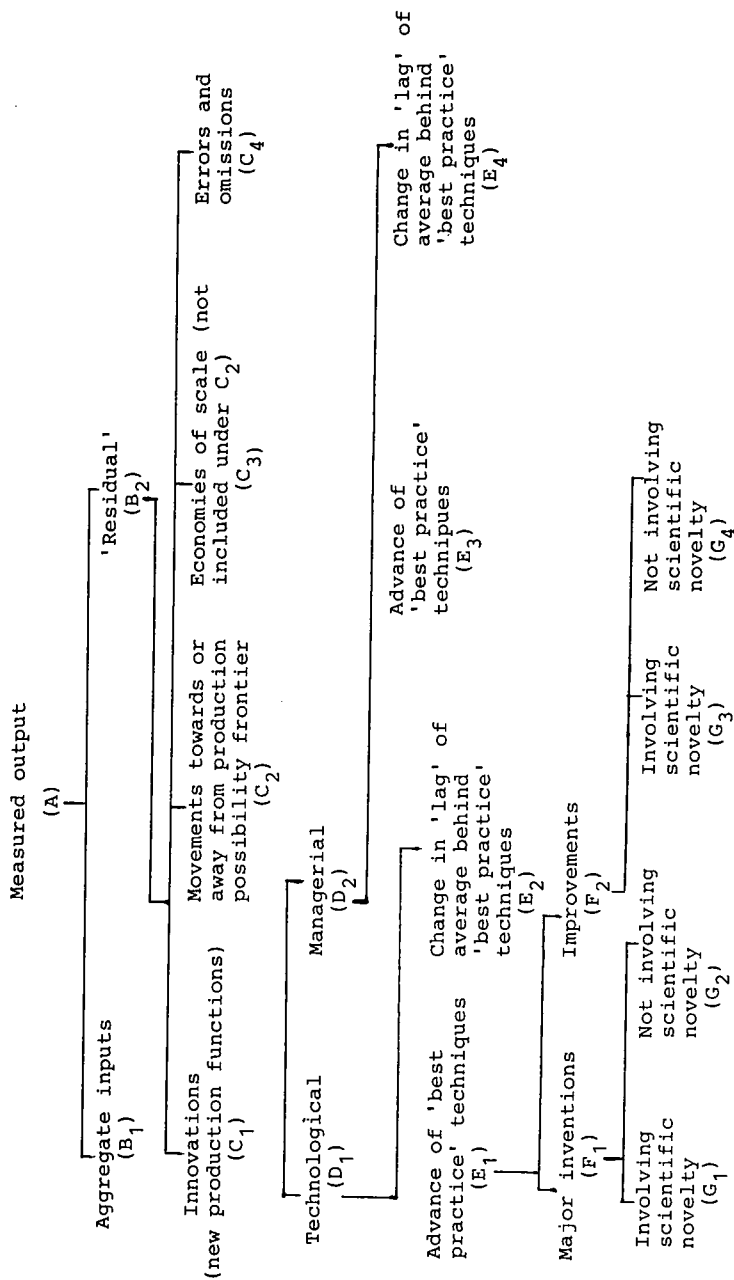
Our point of departure is the discussion of the so-called "residual", the increase in *total* factor productivity which—according to the earliest estimates of aggregate production functions—accounted for between 50 and 85 percent of modern growth in industrial countries such as the U. S. A. (in the 20th century).⁴⁴ Diagram 1 taken from J. D. Gould's excellent survey of growth history, offers an incomplete list of dangers associated with residual analysis. Missing there, but worthy of mention, are quality changes or differences in the inputs of labor (obvious and often noted), capital (less obvious, but also of probable significance), and natural resources. This deserves mention because the productivity of these factors should be standardized according to quality or comparisons will be misleading. In the studies cited by Gould (and others), for example, stronger, healthier and better educated labor as well as improved capital equipment contribute to enlargement of the measured productive inputs and reduction of their measured productivity.⁴⁵ What the diagram—and the literature supporting it—makes clear is that many factors explain (measured) productivity growth. That could mean that the traditional emphasis upon the role of technical change—on one interpretation "D₁" in the diagram and according to another interpretation, "C₁"—requires modification. In its place we might want to install improved allocation of resources (C₂), economies of scale (C₃), and/or what historians tend to call "entrepreneurship" (here D₂). However, no consensus on the weighting of these various components has yet emerged, so it is too early to forecast a justified neglect of technology.⁴⁶

Some of the components listed in Gould's diagram have their counterparts in the recent historical literature on economic growth. Interestingly, agricultural history is relatively rich in examples. To mention just three, David and Griliches for the U.S. in two different periods, and O'Brien and Keyder in their 19th-century British-French comparisons, have identified both improved resource allocation and econom-

Economic Growth, p. 22. However, these results depends on the assumption that the factor inputs have been correctly measured—and the assumption has been questioned. Index number bias may be significant. On this see Usher, *Measurement*, Chapter 12.

44. The earliest and perhaps most striking result was Solow's finding that about 85 per cent of American productivity change, 1909–1949 was attributable to technical change. Correction for the quality of labor, however, reduced this contribution to the share indicated in the previous footnote. Further research brought further fluctuations in the share, one study (by Jorgenson and Griliches) even virtually eliminating technical change (or factor productivity). But see on this Usher, *Measurement*, Chapter 12.
45. In *Why Growth Rates Differ* (Chapter 7–9 and Appendix F) Denison justifies an increase in the growth share attributable to the labor input for "Northwest Europe", 1950–62, from .52 to .83 per cent when the latter is corrected for education and age-sex composition. This reduces the "residual" from 3.33 to 3.07 in percentage points of change per annum.
46. Much depends on how technical change has been and is understood. One can clearly forecast a reduced importance for concern with technical change in the narrow, engineering sense. Technical change, however, can take on an extremely broad meaning, e.g., to include changes in the quality of output. It is not inconceivable that experiments along these lines could enlarge the residual once again. On this see Usher, *Measurement*, Chapter 12.

Diagram 1



The sources of increase in (measured) output

Source: J.D. Gould, *Economic Growth in History*, P. 299

ies of scale as major sources of productivity growth and/or productivity differences.⁴⁷ D. Mc Closkey, R. Allen, and S. Webb, on the other hand, have discussed these sources plus entrepreneurship and technical change against the background of the comparative history of heavy industrial growth in Great Britain, Germany and the U.S. in the 19th century—without arriving at any general agreement on their relative importance.⁴⁸ Disagreement can be instructive, however, and in the case at hand, we come to realize how much results depend upon assumptions (e.g., concerning the degree of competition), the choice of data, and/or the scope of the investigation. For example, whereas the Mc Closkey study concluded—on the basis of international input and output price comparison and the competitive, marginal-cost pricing assumption—that British steel producers were losing ground to U.S. and German rivals towards the end of the 19th century because of demand and relative input cost shifts but *not* because of declining relative efficiency, both Allen and Webb challenged this pricing assumption—particularly as applied to German data. Allen examined both German and American costs and prices and Webb German ones, and both showed that the use of cost data not deduced from prices indicates a clear British lag in relative efficiency, i.e. a lag in relative total factor productivity. What is not clear from all of this, as suggested above, is whether such a lag represents a deficiency in “entrepreneurship”, a technological gap, an economies of scale gap, a lack of demand, or all of these things. The answer to this question, it seems, will require further modelling.⁴⁹

It would be easy to expand on the relationship between the “residual” and comparative productivity history. The preceding remarks have barely scratched the methodological surface. But instead of continuing this rambling journey through the litera-

47. David P., *The Mechanization of Reaping in the ante-bellum Midwest*, reprinted in David P., *Technical Choice, Innovation and Economic Growth, Essays on American and British Experience in the Nineteenth Century*, Cambridge 1975; Griliches Z., *The Sources of Measured Productivity Growth: United States Agriculture, 1940–60*, in: *Journal of Political Economy*, 71 (1963); O'Brien and Keyder, *Economic Growth*, Chapter 5. It may be worth adding that this is consistent with the traditional view that the reallocation of labor out of agriculture into industry must have been a major source of aggregate productivity growth in the 19th century, but that consistency cannot be assumed. For the postwar European situation see Gould, *Economic Growth*, p. 320 (drawing on Denison, *Why Growth Rates Differ*).

48. McCloskey D., *Economic Maturity and Entrepreneurial Decline: British Iron and Steel, 1870–1913*, Cambridge 1973; Allen, *International Competition*; Webb, *Tariffs*.

49. The summary in the text cannot pretend to do justice to a complicated “cliometric” debate. On the German side, Webb stresses market imperfection and scale economies leading to more investment and “embodied” technical change. By implication, the lack of these factors could explain Britain’s relative stagnation. For Britain, however, McCloskey and even Allen acknowledge natural resource disadvantages with consequences for technical change, profitability and investment. McCloskey, to be sure, tends to stress slow demand in Britain rather than slow productivity as the reason for this industry’s relative decline and he expressly rejects the thesis of “entrepreneurial failure”, whereas Allen finds slow productivity change and possibly faulty entrepreneurship. Given the comparative nature of much of the analysis, it is rather surprising to find virtually no discussion of index number problems—either on the output or input side. That could be a source of bias in the results reported. On the whole, this discussion casts doubt on the operational usefulness of the residual components of Diagram 1 for quantitative sectoral studies of productivity.

ture, let me conclude by returning, once more, to the discussion of Fogel and Engerman's controversial interpretation of slave agriculture in the ante-bellum American South.

The justification for taking up this particular example is that it illustrates better than most the key significance which productivity measurement can have for major historiographical questions. I believe that Fogel and Engermann approached an important question—the problem of southern slavery—correctly. They wished to establish, in a quantitative manner, what the economic bases of southern slavery were, without losing sight of the broader issues of social and political history at stake. I happen to disagree with their own interpretation of their findings,⁵⁰ but for present purposes it is the findings themselves and their derivations which deserve discussion. It must suffice here to point out that both findings and interpretations have enormous implications for history, that is, that the example is not a trivial, obscure, or narrowly technical one.

Our concern is with the productivity or “relative efficiency” question. Fogel and Engerman argued that, contrary to what a generation of historians had believed and propagated, the ante-bellum slave economy was not only highly profitable for its plantation-owning elite but also relatively efficient—as measured by its aggregate rate of growth of output per head of population and per input of productive factors, and as measured by the level of its output per factor input in relation to the rest of the United States. By a process of elimination, those authors attributed this “relative efficiency”—and here is where controversy arose—to the high quality or efficiency of slave labor. Our discussion touches on four points: (1) the representativeness of the data; (2) the identification of labor efficiency as the key factor; (3) the quality of labor inputs; and (4) the quality of land inputs. (1) Despite its lengthy discussion of sources and methods (in Volume II) *Time on the Cross* does not permit readers to clearly identify the sample of evidence going into the estimates of relative efficiency (or total factor productivity). It is clear that the U.S. Census of Agriculture for 1860, with data on individual farms and plantations operating in that year, is the main source. We also know the basic sample used (taken from the 1860 census and known as the “Parker-Gallmann sample” after the names of the two economic historians who produced it). However, that information on the size distribution of these production units or on the variance of productivity measures across it which *Time on the Cross* gives does not suffice to firmly establish the representativeness of the data

50. I disagree with their clear distinction between slavery as a retrograde political and moral institution and slave agriculture as a viable, successful economic system. Since the essence of slavery was, in G. Wright's succinct words, the “involuntary reallocation of family labor from nonmarket economic activity to production of crops for sale,” it meant a loss of leisure and freedom which had, in terms of the imputations discussed earlier, a significant economic dimension imperfectly reflected in the market results summarized under the heading of “efficiency”. Cf. Wright G., *The Efficiency of Slavery: Another Interpretation*, in: *American Economic Review*, 69 (1979), p. 225. I also disagree with their belief that the demonstration of high productivity and positive work attitudes among black slaves rescues them from posthumous denigration. On the contrary, one could argue that since, historically, much meaningful labor protest has taken the form of sabotage of work processes through slowdowns, absenteeism and breakage, Fogel and Engerman's attempt to find positive work attitudes among slaves amounts to an attempt to denigrate them.

base. Since the basic, Parker-Gallmann sample was restricted to counties specializing in cotton production, it is hard to know how generalizations about the South as a whole or comparison between free farming and slave farming within the South should be assessed. Thus, the findings that slave agriculture was 28 percent more efficient than free agriculture in the South in 1860 or that Southern agriculture was 35 percent more efficient than Northern agriculture in the same year are difficult to interpret.⁵¹ Finally, we are not told to what extent the findings are sensitive to the disproportionate weight of 1860 data. It seems doubtful whether the entire ante-bellum period (1820–60?) is well characterized by this procedure. (2) Among the many components of productivity increases, *Time on the Cross* singles out the “personal labor efficiency” of black slaves as the main source of the superiority of Southern slave agriculture (in Diagram 1, C 4). In so doing, they assign to the other components such as economies of scale, technical know-how or managerial ability, an insignificant role. They do find scale effects on large, slave plantations, but choose to interpret them in part as a product of the character of slave labor. Management ability receives the same treatment: “In a certain sense”, they write, “all, or nearly all of the advantage is attributable to the high quality of labor, for the main thrust of management was directed at improving the quality of labor”.⁵² The trouble is, there is no clear demonstration of how their implicit weighting of the sources of productivity differentials (between slave and free labor) can be justified. (3) Fogel and Engerman’s comparative estimates of agricultural labor inputs in 1860 are biased upward against the North (i. e., they understate the Southern input relative to the North). The reason for this is that they treat a southern man-year as equal to a northern man-year in spite of the fact that (a) the southern climate permitted a fuller utilization of the entire work-year—possibly by as much as 60 days, or at least one third of the northern work-year—and although (b) slaves worked more hours per year than did free farm workers (in the North and South) and were significant, of course, only in the South.⁵³ Plausible corrections, suggested by David and Temin in their critique of *Time on the Cross* and based on a comparison of hours worked by blacks before and after eman-

51. The two estimates may not be comparable since the latter figures must represent a weighted average of a sample whereas the former may or may not. They rest on the Parker-Gallman sample, but nothing is said on extrapolation techniques or sample variances. In fact, the book contains very little raw data with which readers could check on the Fogel-Engerman claims. See David and Temin, *Progressive Institution?*, pp. 764–65, esp. notes 37 and 38.

52. Fogel and Engerman, *Time on the Cross*, I, p. 210. I find this plausible, given the nature of slave agriculture and the input shares Fogel and Engerman use. However, those shares are not explicitly justified—they could embody an index number problem of their own—and even if they could be deemed satisfactory, we are given little basis for choosing between a high work intensity caused by management and that produced voluntarily by labor. See Hasckell T., *Explaining the Relative Efficiency of Slave Agriculture in the Antebellum South: A Reply to Fogel-Engerman*, in: *American Economic Review*, 69 (1979).

53. As David and Temin point out, *Time on the Cross* does not document its assertion concerning the rough equality of the workyear in slave and free agriculture. Moreover, in one huge correction they deduct 25 per cent of rural slave labor from their input measure on the grounds that this share represented “domestics”. But they do not make it clear whether this deduction is maintained in the next estimate, which is corrected (downward) for age-earnings profiles. This amounts to a bias against the North whose labor force is not so corrected.

cipation, lead to an increased labor input to slave agriculture of between 28 and 34 percent. (4) By using average land values to adjust acreage figures and, hence, land inputs in North and South, *Time on the Cross* implicitly applies the theory of rent to natural differences in soil fertility and assumes perfectly competitive national capital, land and agricultural product markets. But given the high rate of interest on mortgages in the South and the preferred position of presumably risk-averse southern landowners in the land market there, land prices in the South might well tend to understate the productive contribution of land's natural fertility. Moreover, land prices would also be likely to be lower in an area relatively poorly served by transportation facilities—such as the South—because prices received by farms for agricultural products in the South would be lower relative to final market prices (according to which outputs were weighted) than those received by farmers in the North. This could be corrected for by reducing southern land productivity by the appropriate amount (or by other means). In any case, recalculation of land inputs along these lines would increase Southern land inputs relative to the North.⁵⁴

The upshot of this brief discussion is that a great deal can depend on historical productivity measurement, whereas the latter can depend, in turn, on assumptions about outputs, inputs, and residual components whose verification may be extremely difficult. In the present example, had the questioning of such assumptions early on led in the directions suggested here, the puzzle of a high-productivity slave economy propelled by its qualitatively superior labor force might have never emerged. The moral of the story, for our purposes, is not simply that a provocative thesis may be essential to get research activity into significant problems going, but also that where such research is dominated by systematic and explicitly quantitatively comparative methods, corrections of provocative theses and puzzles are possible.

Conclusion

Given the essentially descriptive and taxonomical character of this paper, it would be inappropriate to conclude with statements even more sweeping and summary than those already made. I have attempted, no doubt presumptuously, to define the field of national income and productivity history from the point of view of methodology. At the same time, the paper seeks to call attention to certain areas within the field which coincide, in my opinion, with the research needs and interests of comparative Europeans economic history (the comparative history of prices and consumption pattern and tastes stands out, it seems to me, as such an area). And finally, there is the question of the wider implications of comparative income and productivity history. The last section of the paper discussed one example taken from American history (though the comparison was *intra*-national rather than international) but European history must also be bristling with comparable issues. One such might be seen in the liberal agrarian reforms carried out in the nineteenth century, and to which one may attach the question: did they help, hinder, or result from productivity change? And to what

54. There is some reason for believing that national price weights overvalue southern agricultural output relative to the North in 1860. Cf. David and Temin, *Progressive Institution?* p. 774.

extent was the rise of free trade in western Europe during the middle decades of the nineteenth century the product of rising income and productivity levels? No doubt, such themes lack the political clout of the problems of American slavery. But what about the German Weimar Republic and its aftermath? If the problem of the Weimar Republic in Germany can be seen as dependent upon gaps between real wages and productivity, how are we to interpret that gap in other European countries during the same time? And what about subsequent periods? Lack of patience, lack of knowledge, and perhaps lack of imagination, on my part, limit the list; but I am sure it can be extended. The more difficult task will be the income and productivity studies themselves. They will be welcome.

Zusammenfassung:

Pro-Kopf-Einkommen und Produktivität als Indikatoren für Entwicklung und Wohlstand. Bemerkungen zur Kuznetsianischen Wirtschaftsgeschichte

1. Die Verwendung der makroökonomischen Größe Volkseinkommen pro Einwohner (VpE) in wirtschaftshistorischen Untersuchungen hängt mit dem sog. „Wachstumsparadigma“ zusammen: geht es in solchen Untersuchungen um die Beschreibung und/oder Interpretation des Wirtschaftswachstums, so ist jene Größe (oder ein gleichwertiger Ersatz) – und deren Komponenten – unentbehrlich. Kritik an das Paradigma sollte dennoch von Kritik an der Größe selbst unterschieden werden.
2. Dieser Beitrag geht von der Annahme aus, daß Länder, wie einzelne Wirtschaftssubjekte, nach materieller Wohlstandsteigerung streben, und daß der Erfolg dieses Strebens, d. h. Wohlfahrt, an der Entwicklung des VpE gemessen werden kann.
3. Kritik an der Verwendung des VpE als Wohlstandsindikator kann in 5 Punkten aufgedeutet werden: (1) das Problem nichtmaterieller Güter; (2) das Problem der nichtmarktwirtschaftlichen Aktivitäten; (3) die Definition der End- bzw. Zwischenprodukte; (4) die Annahme von konstanten Präferenzen (einschließlich Indexzifferprobleme); und (5) das Problem der Verteilung. Diese Probleme qualifizieren die Interpretation von VpE als Indikator der Wohlfahrtssteigerung, rechtfertigen aber nicht deren gänzliche Zurückweisung.
4. Alternativen zum VpE wie z. B. Tobin und Nordhaus' M. E. W. oder die „sozialen Indikatoren“ sind in einzelnen Fällen empfehlenswert aber kein allgemein akzeptabler Ersatz für das VpE.
5. Produktivitätsmessung als Instrument der wachstumsorientierten Geschichtsschreibung unterliegt demselben methodischen Problem wie das VpE. Dies kann an der Diskussion einiger neuerer wirtschaftshistorischer Arbeiten insbesondere der Arbeit von Fogel und Engerman über die amerikanische Sklavenwirtschaft im 19. Jahrhundert, gezeigt werden. Diese Diskussion zeigt aber zugleich die große Bedeutung der „Produktivitätsgeschichte“ für allgemeine historiographische Fragen.